I. AMENDMENTS TO THE SPECIFICATION

Replace the paragraph on page 8, line 21 to page 9, line 10 with:

The primers herein are selected to be "substantially" complementary to the different strands of each specific sequence to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands. Therefore, the primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the strand to be amplified to hybridize therewith and thereby form a template for synthesis of the extension product of the other primer. Computer generated searches using programs such as Primer3 (available at genome.wi.mit.edu/cgi-bin/primer/primer3.cgi ~~www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi~~), STSPipeline (available at genome.wi.mit.edu/cgi-bin/www-STS_Pipeline ~~www-genome.wi.mit.edu/cgi-bin/www-STS_Pipeline~~) or GeneUp (Pesole *et al.*, *BioTechniques* *25*:112-123 (1998)) for example, can be used to identify potential PCR primers. Exemplary primers include primers that are 18 to 50 bases long, where at least between 18 to 25 bases are identical or complementary to at least 18 to 25 bases segment of the template sequence. Preferred template sequences for such primers are selected from a fragment of any one of SEQ ID NO: 138061 through SEQ ID NO: 195836 or complements thereof.

Replace the paragraph on page 36, lines 8 to 16 with:

PHRED is used to call the bases from the sequence trace files (available at mbt.washington.edu ~~http://www.mbt.washington.edu~~). PHRED uses Fourier methods to examime the four base traces in the region surrounding each point in the data set in order to predict a series of evenly spaced predicted locations. That is, it determines where the peaks would be centered if there were no compressions, dropouts or other factors shifting the peaks from their "true" locations. Next, PHRED examines each trace to find the centers of the actual, or observed peaks and the areas of these peaks relative to their neighbors. The peaks are detected independently along each of the four traces so many peaks overlap. A dynamic programming algorithm is used to match the observed peaks detected in the second step with the predicted peak locations found in the first step.

Replace the paragraph on page 36, line 23 to page 37, line 11 with:

Contigs are assembled using PANGEA clustering tools (PANGEA SYSTEMS, INC. ~~SYSTEMS, INC~~) and PHRAP (available at mbt.washington.edu ~~http://www.mbt.washington. edu~~). PANGEA clustering tools are a series of scripts which group sequences (clusters) by comparing pairs of sequences for overlapping bases. The overlap is determined using the following high stringency parameters: word size = 8; window size = 60; and identity is 93%. Each of the clusters are then assembled using PHRAP. This step results in 102,349 islands. The next step is to combine the islands together to collapse the contig number even further. Default, less stringent parameters, are used in this step: minimum match = 14, minimum score = 30; and the penalty is –2. The resulting, final, total number of islands equals 81,306. Of these, 50,262

are actual contigs and 31,044 are singletons. The final set of 81,306 nucleic acid sequences is identified in Table 1 by sequence identification (Seq. ID) "ATL8Sxxxxx" (for singleton sequences) and "ATL8Cxxxxx" (for contig sequences) and by the corresponding sequence number (SEQ ID NO: 1 through SEQ ID NO: 81,306). The final set of 81,306 genomic sequences is run through the following annotation and gene selection processes described in Example 4. The genomic sequence traces and many of the contigs and singleton traces are disclosed in copending provisional applications for patent identified by serial nos. 60/111,990; 60/111,991; 60/114,151; 60/120,644; 60/135,825; 60/139,932; 60/143,994 and 60/155,422.